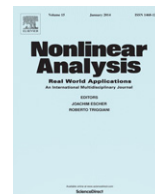


Contents lists available at [ScienceDirect](http://ScienceDirect.com)

Nonlinear Analysis: Real World Applications

journal homepage: www.elsevier.com/locate/nonrwa

Nonlinear stability of flock solutions in second-order swarming models



J.A. Carrillo, Y. Huang, S. Martin*

Department of Mathematics, Imperial College London, London, SW7 2AZ, UK

ARTICLE INFO

Article history:

Received 19 September 2013

Accepted 9 December 2013

ABSTRACT

In this paper we consider interacting particle systems which are frequently used to model collective behaviour in animal swarms and other applications. We study the stability of orientationally aligned formations called flock solutions, one of the typical patterns emerging from such dynamics. We provide an analysis showing that the nonlinear stability of flocks in second-order models entirely depends on the linear stability of the first-order aggregation equation. Flocks are shown to be nonlinearly stable as a *family* of states under reasonable assumptions on the interaction potential. Furthermore, it is tested numerically that commonly used potentials satisfy these hypotheses and the nonlinear stability of flocks is investigated by an extensive case study of uniform perturbations.

© 2014 The Authors. Published by Elsevier Ltd. Open access under [CC BY license](http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Self-organization and pattern formation are ubiquitous in nature and science, ranging from animal aggregation [1–5] and biological systems [6,7] to self-assembly of nano-particles [8,9]. The intense studies during the last two decades are on both individual-based systems and continuum equations [10–13]. One of the essential features in these models is the non-zero characteristic speed of the individual agents, which has been modelled in different ways. The speed can be assumed to be constant with a direction based on the averages of the neighbours [14] or to be driven by random noise [15]. On the other hand, a large class of models consist of self-propelled particles powered by biological or chemical mechanisms with friction forces, resulting in a preferred characteristic speed.

In general, the particles with non-zero equilibrium speed do not form any recognizable patterns [4,16], and interactions within the group have to be included to generate interesting spatial configurations. Most of these interaction forces have been taken into account in the combination of three effects: alignment, repulsion, and attraction; also called the “first principles of swarming”. The basic mechanisms account for collisional avoidance and comfort regions (repulsion), grouping and socialization (attraction), and mimetic synchronization (alignment). The combination of these three effects goes back to fish school modelling [17,18] and these basic ideas have been improved and applied to several animal species including different mechanisms and interactions more adapted to particular living organisms, see for instance [19–23].

In this paper, we focus on the cases with velocity-independent interactions. More precisely, by introducing a pairwise symmetric interaction potential $W(x) = U(|x|)$, we consider the two-dimensional model

* Corresponding author. Tel.: +44 2075948396.

E-mail addresses: carrillo@imperial.ac.uk (J.A. Carrillo), yanghong.huang@imperial.ac.uk (Y. Huang), stephan.martin@imperial.ac.uk (S. Martin).

$$\begin{aligned} \frac{dx_i}{dt} &= v_i, \\ \frac{dv_i}{dt} &= \alpha v_i - \beta v_i |v_i|^2 - \sum_{j \neq i} \nabla W(x_i - x_j), \end{aligned} \tag{1}$$

where $x_i, v_i \in \mathbb{R}^2, i = 1, \dots, N$ are the positions and velocities of the individual particles and α, β are effective values for self-propulsion and friction forces, see [10–12] for more discussions.

For this relatively simple system, a variety of patterns are observed, for instance coherent moving flocks and rotating mills [10]. Other patterns like rings and clumps were discovered in [11,12], with the introduction of the concept of *H-stability* from statistical physics to better characterize the patterns in the parameter space. Delta rings, with mass uniformly distributed on a circle, have been further studied thoroughly for power-law like potentials, we refer to [24–30] for details. While the stability and bifurcation of the ring solutions can be investigated in a straightforward way because of their explicit particle representations on a circle, there are few studies on the more prevalent compact steady solutions, like flocks or mills. The reason lies in the difficulties to solve some complicated integro-differential equations for most of the popular choices of the potential W . Only for certain particular potentials like the quasi-Morse type proposed in [31], the existence and uniqueness of the coherent moving flocks can be established rigorously [32].

In this paper, we focus on the stability of flock solutions of (1) for general potentials defined below.

Definition 1. A flock solution of the particle model (1) is a spatial configuration \hat{x} with zero net interaction force on every particle, that translates at a uniform velocity $m_0 \in \mathbb{R}^2$ with $|m_0| = \sqrt{\frac{\alpha}{\beta}}$, hence $(x_i(t), v_i(t)) = (\hat{x}_i - tm_0, m_0)$.

We note, that the spatial configuration \hat{x} is a stationary state to the first-order interacting particle system

$$\frac{dx_i}{dt} = - \sum_{j \neq i} \nabla W(x_i - x_j), \tag{2}$$

which is analyzed for example in [8,9]. In this work, we focus on general flock solutions whose spatial configurations tend to a compactly supported particle density in the continuum limit.

One conclusion from our analysis is the somehow unexpected deep relation between the linear stability of the flock spatial configuration as steady state for the first-order swarming model (2) and the nonlinear stability of the family of associated flock solutions for the second-order swarming model (1). This relation was already found in the linear stability analysis around flock solutions in [30]. There the authors showed that the linearization of (2) around the equilibrium state \hat{x} has a positive eigenvalue if and only if the linearization of (1) around the steady flock solution in the co-moving frame has an eigenvalue with positive real part. Moreover, they show that if the equilibrium state \hat{x} is linearly stable for (2), then the associated flock solution is always linearly unstable due to the presence of a generalized eigenvector associated to the zero eigenvalue of the linearization of the flock solution in the co-moving frame due to symmetries.

However, it is in this work where we clarify completely this deep imbrication by showing the nonlinear stability of the flock solutions under mild hypotheses on the linearization of (2) about the spatial configuration \hat{x} . Actually, the basic hypotheses, apart from few technical assumptions, are that the steady state \hat{x} of (2) is linearly asymptotically stable except the obvious symmetries: translations and rotations. Therefore, we can apply our theorem to verify the nonlinear stability of flock solutions to more biologically adapted potentials such as the Morse and the Quasi-Morse potentials [11].

The main result of this paper asserts, except few technicalities, that *the family of flock solutions* associated to a linearly asymptotically stable steady state of (2) is *asymptotically stable* for the dynamics of (1). Here, the asymptotic stability of the family of flock solutions means that any small enough perturbation in (x, v) -space at any time t_0 will, under the dynamics of the system (1), relax towards (likely) another flock solution in the family at an exponential rate as $t \rightarrow \infty$. Let us finally emphasize that the most rigorous way of stating our main theorem uses advanced concepts of dynamical systems [33]. Our main theorem can be rephrased as follows: the family of flock solutions to (1) associated to a linearly (except symmetries) asymptotically stable steady state of (2) forms a normally hyperbolic invariant manifold for the system (1) with an empty unstable manifold.

This result has important applications, especially in the study of flock patterns of (1) using particle simulations. In general, the desired patterns are observed only after certain transit dynamics, and with carefully prepared initial data. With the spatial configurations from the first order system (2) (if they exist), the computational load is reduced significantly, making possible to study the existence of these patterns in great details.

After some preliminary notations and assumptions, the main result of this paper is stated in Section 2. The formulation of the linearized dynamics when the mean velocity is transformed to zero is given in Section 3, where the non-trivial relation between the Jacobian matrices of the two systems (1) and (2) is clarified. A comparison of this work to the techniques used in [30] also is given. In Section 4, we use the eigenspace structures of these Jacobian matrices to reveal the connection between the nonlinear stability of the flock solutions for (1) and the linear stability of their spatial configuration for (2), leading to the proof of the main theorem. Section 5 is devoted to numerical experiments to check the validity of the assumptions on the linear stability for the steady state of the system (2) for different biologically reasonable potentials.

2. Main result

In this section, we formulate the main theorem of this work. We will start with some assumptions on the spatial configuration of flock solutions, which is a stationary state of the first-order swarming model (2). The hypotheses assume this spatial configuration to be stable under the dynamics of (2).

Given a steady spatial particle configuration to (2), i.e. a set of particle positions $(\hat{x}_i)_{i=1}^N$ such that

$$\sum_{j \neq i} \nabla W(\hat{x}_i - \hat{x}_j) = 0, \quad \text{for all } i \in \{1, \dots, N\}, \quad (3)$$

we denote by \hat{x} the $2N$ -dimensional spatial configuration vector $(\hat{x}_1, \dots, \hat{x}_N)$. We can now write the linearization of the system (2) at the steady configuration \hat{x} as

$$\frac{dh}{dt} = G(\hat{x})h, \quad (4)$$

where $G(\hat{x})$ denotes the Jacobian matrix associated to (2) and $h \in \mathbb{R}^{2N}$. The $2N \times 2N$ -Jacobian matrix is explicitly given by $G(\hat{x}) = (G_{ij})$ with G_{ij} being the 2×2 -blocks defined as

$$G_{ij} = \begin{cases} -\sum_{j \neq i} \text{Hess } W(\hat{x}_i - \hat{x}_j) & \text{for } i = j \\ \text{Hess } W(\hat{x}_i - \hat{x}_j) & \text{for } i \neq j, \end{cases} \quad (5)$$

with $\text{Hess } W$ denoting the Hessian matrix of W .

We will make extensive use of the following standard notation: given two matrices $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{p \times q}$, the Kronecker product $A \otimes B$ is defined as the matrix $A \otimes B = (a_{ij}B)_{ij} \in \mathbb{R}^{np \times mq}$. Let us also denote by 0_n , 1_n the column vectors of length n with all entries equal to 0 or 1 respectively, and I_n the identity matrix of size n . Let us now specify the precise assumptions on \hat{x} , and thus indirectly on the potential W .

Assumption 1. Given an interaction potential W , we assume the existence of a linearly asymptotically stable (except symmetries) stationary state not lying on a line, i.e., a spatial configuration \hat{x} such that the following assumptions hold:

- (H1) \hat{x} is a stationary state of (2), i.e. \hat{x} satisfies (3).
- (H2) The eigenspace for the zero eigenvalue of $G(\hat{x})$ is spanned by the eigenvectors

$$w_1 = 1_N \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad w_2 = 1_N \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad w_3 = \left(I_N \otimes \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right) \hat{x} \quad (6)$$

representing invariance of $G(\hat{x})$ with respect to translations and rotations in \mathbb{R}^2 , and therefore is three dimensional, $\dim(\text{Eig}(G(\hat{x}), 0)) = 3$.

- (H3) All other eigenvalues of $G(\hat{x})$ have negative real parts, i.e. $\text{Re}(\lambda_i) < 0 \Leftrightarrow \lambda_i \neq 0$ for all $i \in \sigma(G(\hat{x}))$.
- (H4) Not all points of \hat{x} lie on a straight line in \mathbb{R}^2 .

Hypotheses (H1)–(H3) imply that \hat{x} is linearly asymptotically stable for (4) except for the obvious symmetries, or equivalently that the Jacobian matrix $G(\hat{x})$ is negative definite except for the null eigenspace in (6). As a consequence, \hat{x} is nonlinearly asymptotically stable for the dynamics of (2) except for translations and rotations. More rigorously speaking and making use of more advanced dynamical systems theory [33], properties (H1)–(H3) imply that the family of stationary states of (2) obtained by rotations and translations from \hat{x} given by

$$RT(\hat{x}) = \{y : y = 1_N \otimes b + (I_N \otimes R(\phi))\hat{x}, \phi \in [0, 2\pi), b \in \mathbb{R}^2\}$$

forms a normally hyperbolic invariant manifold for the system of ODEs in (2) for which the unstable manifold is empty. Here,

$$R(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

is the 2×2 rotation matrix of angle ϕ . This stability concept will be further discussed and used later on for the second order model (1).

Let us remark, that assumptions (H1)–(H4) are very mild in the sense that they are natural properties of stationary states of (2) found through particle simulations from generic initial configurations, e.g. compactly supported flocks. Particular special cases satisfying hypotheses (H1)–(H4) are the so-called Delta rings: particles arranged equidistantly on a circle. We will discuss in the last section more biologically relevant potentials, such as Morse-like potentials [8–11,31,32], for which we can numerically find stationary states \hat{x} satisfying hypotheses (H1)–(H4).

Now, let us make precise the concept of the family of flock solutions of (1) associated to the stationary state \hat{x} of (2). Let us denote by \mathcal{V}_s the set of possible steady velocity vectors for a flock solution of speed $s = \sqrt{\frac{\alpha}{\beta}}$ to (1) defined as

$\mathcal{V}_s = \{m_0 \in \mathbb{R}^2 : |m_0| = s\}$. Then the family of flock solutions of (1) associated to the steady state \hat{x} of (2) is given by

$$\mathcal{Z}_F = \left\{ z = \begin{pmatrix} x^* + v^*t \\ v^* \end{pmatrix}_{t \in \mathbb{R}}, v^* = 1_N \otimes m_0, x^* \in RT(\hat{x}), m_0 \in \mathcal{V}_s \right\}. \tag{7}$$

It is straightforward to check that all $z \in \mathcal{Z}_F$ are solutions to (1) in vectorial form. The next section will be devoted to analyse the system (1) in suitable coordinates to study their stability. A change of variables to (1) that eliminates the translation of one particular member of the flock family is given by the co-moving frame $\tilde{v}_i := v_i - m_0$ transforming the flock $(x^* + v^*t, v^*)$ onto a stationary state. However, it is very natural to think of small perturbations on $(x^* + t v^*, v^*)$ that, under the dynamics of (1), tend to possibly another spatial configuration in $RT(\hat{x})$ translating at a different velocity generated by \tilde{m}_0 , i.e. to another member of the family of flocks \mathcal{Z}_F . Hence, we cannot expect linear stability for general perturbations for such a change of variables... Instead, we propose a new change of variables measuring the difference to mean velocity in velocity space, and aim to establish nonlinear stability in the entire family \mathcal{Z}_F .

Definition 2. Defining the mean velocity of N particles as

$$m(t) := \frac{1}{N} \sum_{i=1}^N v_i(t),$$

we introduce the change of variables with respect to the mean velocity $m(t)$ as:

$$\begin{aligned} \tilde{x}_i(t) &:= x_i(t) - \int_0^t m(s) ds, \\ \tilde{v}_i(t) &:= v_i(t) - m(t). \end{aligned} \tag{8}$$

We will work with the dynamical system (1) and its linearization in vector form, and thus we need to introduce some matrix notation used in the following.

Definition 3 (Notation).

1. $0_{n \times p}$ is a block of zeros of size $n \times p$.
2. Writing $\delta_n^i \in \mathbb{R}^n$, we refer to a vector of length n with value 1 as its i -th entry and zeros elsewhere.
3. For the sake of simplicity, we occasionally write column vectors as row vectors neglecting the transpose operation on its components, e.g.

$$(x, v, 0_2)^T := (x^T, v^T, 0_2^T)^T = \begin{pmatrix} x \\ v \\ 0_2 \end{pmatrix}.$$

4. For a square matrix $A \in \mathbb{R}^{n \times p}$ or a column vector $b \in \mathbb{R}^n$, we denote by

$$[A] := (a_{ij})_{\substack{i=1, \dots, n-2 \\ j=1, \dots, p}}, \quad [b] = (b_i)_{i=1, \dots, n-2},$$

the matrix obtained from A by eliminating the last two rows and analogously for vectors. The k -th row or column of A is labelled as $a_{k,\cdot}$ or $a_{\cdot,k}$ respectively.

Lemma 1. Using the change of coordinates (8), the microscopic model (1) reads

$$\begin{aligned} \frac{d\tilde{x}_i}{dt} &= \tilde{v}_i, \\ \frac{d\tilde{v}_i}{dt} &= \frac{N-1}{N} (\alpha - \beta |\tilde{v}_i + m|^2) (\tilde{v}_i + m) - \frac{1}{N} \sum_{j \neq i} (\alpha - \beta |\tilde{v}_j + m|^2) (\tilde{v}_j + m) - \sum_{j \neq i} \nabla W(\tilde{x}_j - \tilde{x}_i), \\ \frac{dm}{dt} &= \frac{1}{N} \sum_j (\alpha - \beta |\tilde{v}_j + m|^2) (\tilde{v}_j + m). \end{aligned} \tag{9}$$

For any mean velocity $m_0 \in \mathcal{V}_s$, the vector $\hat{Q} = (\hat{x}, 0_{2N}, m_0)^T$ is a stationary state of (9). Moreover, the Jacobian matrix F obtained by linearizing (9) around a stationary state \hat{Q} with an arbitrary orientation m_0 is given by

$$F = \begin{pmatrix} 0_{2N \times 2N} & & I_{2N} & & 0_{2N \times 2} \\ G(\hat{x}) & - \begin{pmatrix} \frac{N-1}{N} & -\frac{1}{N} & \cdots & -\frac{1}{N} \\ -\frac{1}{N} & \ddots & & \vdots \\ \vdots & & \ddots & -\frac{1}{N} \\ -\frac{1}{N} & \cdots & -\frac{1}{N} & \frac{N-1}{N} \end{pmatrix} & \otimes 2\beta(m_0 \otimes m_0^T) & 0_{2N \times 2} \\ 0_{2 \times 2N} & & -1_N^T \otimes \frac{2\beta}{N}(m_0 \otimes m_0^T) & & -2\beta(m_0 \otimes m_0^T) \end{pmatrix}. \tag{10}$$

Proof. By definition $\frac{d\tilde{x}_i}{dt} = \frac{dx_i}{dt} - m(t) = v_i(t) - m(t) = \tilde{v}_i(t)$. Next we note $\tilde{x}_i - \tilde{x}_j = x_i - x_j$, and hence $\nabla W(\tilde{x}_i - \tilde{x}_j) = \nabla W(x_i - x_j)$. Now, we compute

$$\begin{aligned} \frac{dm}{dt} &= \frac{1}{N} \sum_i \frac{dv_i}{dt} = \frac{1}{N} \left(\sum_i (\alpha - \beta|v_i|^2) v_i + \sum_i \sum_{j \neq i} \nabla W(x_i - x_j) \right) \\ &= \frac{1}{N} \sum_i (\alpha - \beta|\tilde{v}_i + m|^2) (\tilde{v}_i + m), \end{aligned}$$

using that the net total force in the system is zero by symmetry, i.e. Newton’s action–reaction principle. From this, the equation for $\frac{d\tilde{v}_i}{dt} = \frac{dv_i}{dt} - \frac{dm}{dt}$ is readily available. Now, any state $\hat{Q} = (\hat{x}, 0_{2N}, m_0)^T$, with $m_0 \in \mathcal{V}_s$ is a stationary state of (9) since $G(\hat{x}) = 0$ and all self-propulsion terms vanish. Let us drop the tildes henceforth for simplicity and let us compute the Jacobian matrix of the system (9). The first line of blocks in F is trivial. In the second line of blocks, we first have $G(\hat{x})$ since the dependency on x in (9) is identical to (1), hence one obtains the Jacobian of the first-order particle system (2) evaluated at the stationary spatial configuration. To compute the second block, consider an arbitrary particle i and its velocity vector $(v_{i,1}, v_{i,2})^T$. Then, identifying the right-hand sides of (9) with the time derivative symbols, we get:

$$\begin{aligned} \partial_{v_{i,1}} \frac{dv_{i,1}}{dt} &= \frac{N-1}{N} (-2\beta(v_{i,1} + m_1)^2 + (\alpha - \beta|v_i + m|^2)), \\ \partial_{v_{i,2}} \frac{dv_{i,1}}{dt} &= \frac{N-1}{N} (-2\beta(v_{i,2} + m_2)(v_{i,1} + m_1) + (\alpha - \beta|v_i + m|^2)), \\ \partial_{v_{j,1}} \frac{dv_{i,1}}{dt} &= -\frac{1}{N} (-2\beta(v_{j,1} + m_1)^2 + (\alpha - \beta|v_j + m|^2)), \end{aligned}$$

for $j \neq i$ and analogous expressions for $\partial_{v_{i,1}} \frac{dv_{i,2}}{dt}$, $\partial_{v_{i,2}} \frac{dv_{i,2}}{dt}$, $\partial_{v_{j,2}} \frac{dv_{i,1}}{dt}$, $\partial_{v_{j,1}} \frac{dv_{i,2}}{dt}$, and $\partial_{v_{j,2}} \frac{dv_{i,2}}{dt}$. Evaluating at \hat{Q} the last expression in all terms vanishes since $|m_0|^2 = \frac{\alpha}{\beta}$, which in total gives the structure of the block using the Kronecker product notation. In the third line of F , the second block follows analogously. For the last block, we compute

$$\partial_{m_1} \frac{dm_1}{dt} = \frac{1}{N} \sum_j (-2\beta(v_{j,1} + m_1)^2 + (\alpha - \beta|v_j + m|^2)),$$

which reduces to $\partial_{m_1} \frac{dm_1}{dt} = -2\beta m_1^2$ at \hat{Q} , and analogously for the remaining partial derivatives. This completes the computation of the structure of F . ■

Let us point out that the family of flock solutions to (1) is translated through the change of variables (8) to the set of stationary solutions of (9) given by

$$\tilde{\mathcal{Z}}_F = \left\{ Q^* = \begin{pmatrix} x^* \\ 0_{2N} \\ m_0 \end{pmatrix}, x^* \in RT(\hat{x}), m_0 \in \mathcal{V}_s \right\}. \tag{11}$$

Now, we have almost all the ingredients to write the main result of this paper that holds under an additional technical assumption on the interplay between the linearization and the orientation of the flock.

Assumption 2 (H5). All eigenvectors to non-zero eigenvalues of $G(x^*)$ are not particle-wise orthogonal to the fixed mean-velocity m_0 . This means, that for pairs of entries $2i - 1, 2i$ of any eigenvector w of $G(x^*)$, we have

$$\left\langle m_0, w_{\begin{smallmatrix} 2i-1 \\ 2i \end{smallmatrix}} \right\rangle \neq 0 \quad \text{for at least one } i \in 1, \dots, N.$$

We will understand the role of this additional assumption below, but let us point out that this assumption is generically met by the members of the flock family $\tilde{\mathcal{Z}}_F$. The main result of this work reads as:

Theorem 1. Assume that the symmetric pairwise interaction potential W allows for a stationary spatial configuration \hat{x} of (2) satisfying assumptions (H1)–(H4). Consider the second-order swarming model (1) with $\alpha, \beta > 0$, then the family of flock solutions \mathcal{Z}_F of (1) defined in (7) is locally asymptotically stable for all configurations x^* satisfying (H5) in the following sense: Any small enough perturbation in (x, v) -space at any time t_0 of the flock solution $z^* \in \mathcal{Z}_F$ associated to x^* will, under the dynamics of the system (1), relax towards another flock solution $\tilde{z} \in \mathcal{Z}_F$ at an exponential rate as $t \rightarrow \infty$.

Let us remind that assumptions (H1)–(H4) imply in particular that the family of stationary states $RT(\hat{x})$ forms a normally hyperbolic invariant manifold for (2) with empty unstable manifold. Therefore, our main theorem can be rephrased as stating that the family of flock solutions associated to \hat{x} of the second order model (1) has essentially the same property. More precisely, we have the following result.

Corollary 1. Assume that the symmetric pairwise interaction potential W allows for a stationary spatial configuration \hat{x} of (2) satisfying assumptions (H1)–(H4). Then the subset of stationary states of $\tilde{\mathcal{Z}}_F$ in (11) satisfying (H5) is a normally hyperbolic invariant manifold of (9) with empty unstable manifold.

Next two sections are devoted to the proof of these results. We start in next section by restricting the dynamics to a physically consistent subspace.

3. Physically consistent dynamics and its linearization

Note that (9) is a $4N + 2$ dimensional system and thus $F \in \mathbb{R}^{(4N+2) \times (4N+2)}$ by having added the mean velocity as an additional variable. However, a stability analysis of (10) would include unphysical perturbations where particle velocities do not match mean velocity. The dynamics we are interested in are $4N$ -dimensional, and thus we need to reduce (10) to physically consistent dynamics. Clearly, mean velocity consistent states are invariant under the dynamics of (9), since

$$\frac{d}{dt} \frac{1}{N} \sum_i (v_i - m) = 0.$$

We hence define a base for the $4N$ -dimensional subspace of mean velocity consistent states by removing the redundant equation of the N -th particle.

Definition 4. Define the subspace of mean velocity consistent states as $\text{span}(B)$ with $B = \{b_i\}_{i=1, \dots, 4N}$ and

$$b_i = \begin{cases} (\delta_{2N}^i, & 0_{2N}, & 0_2)^T & i = 1, \dots, 2N, \\ (0_{2N}, & \delta_{2N-2}^k & -1 & 0, & 0_2)^T & i = 2N + (2k - 1), k = 1, \dots, N - 1, \\ (0_{2N}, & \delta_{2N-2}^k & 0 & -1, & 0_2)^T & i = 2N + 2k, k = 1, \dots, N - 1, \\ (0_{2N}, & 0_{2N}, & \delta_2^k)^T & i = 4N - 2 + k, k = 1, 2. \end{cases}$$

Any physically relevant perturbation of (1) is projected onto $\text{span}(B)$ via

$$P : \begin{pmatrix} \Delta x \\ \Delta v \end{pmatrix} \mapsto \begin{pmatrix} \Delta x \\ \text{---} \\ \Delta v_1 - \frac{1}{N} \sum \Delta v_i \\ \vdots \\ \Delta v_N - \frac{1}{N} \sum \Delta v_i \\ \text{---} \\ \frac{1}{N} \sum \Delta v_i \end{pmatrix} \in \text{span}(B).$$

Clearly, B is isomorphic to a base of \mathbb{R}^{4N} and the relevant invariant subspace for the dynamics of (9). As mean velocity consistency is a linear algebraic constraint to (9), it turns out that $\text{span}(B)$ is also an invariant subspace of F , which means we can restrict the stability analysis of (10) to physically relevant perturbations by computing the matrix representation of $F : \text{span}(B) \rightarrow \text{span}(B)$ under a change of base to B , denoted by F_B^B .

Lemma 2. $\text{span}(B)$ is an invariant subspace of F . Restricting F to $\text{span}(B)$ and changing base to B gives

$$F_B^B = \begin{pmatrix} 0_{2N \times 2N} & I_{2N-2} & 0_{2N \times 2} \\ [G(\hat{x})] & -I_{N-1}^T \otimes I_2 & 0_{2N-2 \times 2} \\ 0_{2 \times 2N} & 2\beta(m_0 \otimes m_0^T) & -2\beta(m_0 \otimes m_0^T) \end{pmatrix}, \tag{12}$$

where $[\cdot]$ is used according to Definition 3.4.

Proof. We compute the image of base vectors b_i under F . First, consider b_i with a non-zero entry in a spatial coordinate ($i = 1, \dots, 2N$). Then $Fb_i = (0_{2N}, g_{\cdot,i}, 0_2)^T$ consists of the i -th column of $G(\hat{x}) = (g_{i,j})_{ij}$. Since the sum of each column of $G(\hat{x})$ in (5) is zero, we can rewrite this in terms of B as

$$(0_{2N}, g_{\cdot,i}, 0_2)^T = \sum_{l=1}^{2N-2} g_{l,i} b_{2N+l},$$

which gives the first block column of (12). Next, consider b_i with a non-zero entry in a velocity component k and assume i to be odd ($i = 2N + (2k - 1), k = 1, \dots, N - 1$). Then we can compute Fb_i to get

$$\left(\begin{array}{c} \delta_{2N-2}^k \\ -1 \\ 0 \\ \cdots \\ - \left(\begin{array}{ccc} \frac{N-1}{N} & \cdots & -\frac{1}{N} \\ \vdots & \ddots & \vdots \\ -\frac{1}{N} & \cdots & \frac{N-1}{N} \end{array} \right) \\ \cdots \\ 0_2 \end{array} \right) \otimes 2\beta(m_0 \otimes m_0^T) \left(\begin{array}{c} \delta_{2N-2}^k \\ -1 \\ 0 \end{array} \right) = \left(\begin{array}{c} \delta_{2N-2}^k \\ -1 \\ 0 \\ \cdots \\ 0_{2(k-1)} \\ -2\beta(m_0 \otimes m_0^T)_{\cdot,1} \\ \cdots \\ 0_{2(N-k)-2} \\ 2\beta(m_0 \otimes m_0^T)_{\cdot,1} \\ \cdots \\ 0_2 \end{array} \right)$$

since, in the second block, the multiplication with $(\delta_{2N-2}^k, -1, 0)^T$ first gives $\frac{N-1}{N} + \frac{1}{N}$ in row k , $-\frac{1}{N} - \frac{N-1}{N}$ in row $2N - 1$ and zero in all other rows, which is then multiplied with $-2\beta(m_0 \otimes m_0^T)$ to select its first column twice (as i is odd), indicated by the subscript $(\cdot)_{\cdot,1}$. This rewrites in terms of B as

$$Fb_i = b_{i-2N} - b_{2N-1} - 2\beta m_{0,1}^2 b_i - 2\beta m_{0,1} m_{0,2} b_{i+1}.$$

Analogously, one obtains for even i ($i = 2N + 2k, k = 1, \dots, N - 1$)

$$Fb_i = b_{i-2N} - b_{2N} - 2\beta m_{0,1} m_{0,2} b_{i-1} - 2\beta m_{0,2}^2 b_i,$$

involving the second column of $-2\beta(m_0 \otimes m_0^T)$. The change of base for vectors of B with non-zero entry in the mean velocity component is trivial. Having shown that all images of base vectors can be expressed in terms of B , we showed the invariant subspace property and obtain the block structure of F_B^B from our computations. ■

We now derive two corollaries on the linear structure of F_B^B , which will be essential to establish our stability result. To do so, we first rewrite (12) as

$$F_B^B =: \left(\begin{array}{cc} H & 0_{4N-2 \times 2} \\ 0_{2 \times 4N-2} & -2\beta(m_0 \otimes m_0^T) \end{array} \right), \tag{13}$$

since the two blocks separate. We first investigate the eigenspace associated to the zero eigenvalue of F_B^B . . .

Lemma 3. Suppose that $G(\hat{x})$ satisfies (H1)–(H4), then F_B^B does not possess a generalized eigenvector for eigenvalue zero.

Proof. Since the lower 2×2 block of F_B^B in (13) can be diagonalized with eigenvalues 0 and -2α , then we restrict our analysis to H . To simplify the notation, we will skip the dependency of the matrix $G(\hat{x})$ on the stationary state in this proof. If $z = (x, v)^T \in \text{Eig}(H, 0)$ is a zero eigenvector, then $v = 0_{2N-2}$ and $x \in \text{Eig}(G, 0)$ by (12). Suppose $(u, w)^T$ is a generalized eigenvector, then

$$H \begin{pmatrix} u \\ w \end{pmatrix} = \begin{pmatrix} x \\ 0_{2N-2} \end{pmatrix}, \quad x \in \text{Eig}(G, 0)m,$$

which is equivalent to

$$w = [x], \tag{14a}$$

$$\sum_{i=1}^N x_{2i-1} = \sum_{i=1}^N x_{2i} = 0, \tag{14b}$$

$$[Gu] - I_{N-1} \otimes 2\beta(m_0 \otimes m_0^T)w = 0_{2N-2}. \tag{14c}$$

Due to the definition of G in (5), we get

$$\sum_{i=1}^N g_{2i-1,\cdot} = \sum_{i=1}^N g_{2i,\cdot} = 0_{2N}^T. \tag{15}$$

Adding two rows using (15), (14a), and (14b), (14c) is equivalent to

$$Gu - I_N \otimes 2\beta(m_0 \otimes m_0^T)x = 0_{2N}. \tag{16}$$

We left-multiply with x and get

$$0 = x^T Gu = x^T (I_N \otimes 2\beta(m_0 \otimes m_0^T))x, \tag{17}$$

since G symmetric and $x \in \text{Eig}(G, 0)$. This implies $x = [a_1 \dots a_N]^T \otimes m^\perp$ for some constants a_i since the right-sided expression is quadratic and hence x has to be particle-wise orthogonal to m_0 for the equation to hold ($m_0^\perp := (-m_{0,2}, m_{0,1})^T$). On the other hand, (14b) implies that $x \in \text{span}(w_3)$, since w_3 , as given in (6), is the only zero eigenvector of G whose component sums can be zero. Hence, all position vectors are tangential to a fixed mean velocity m_0 , which is a contradiction to (H4). ■

... Next, we investigate the eigenvalues of F_B^B under another hypothesis linking the eigenvectors of $G(\hat{x})$ with the orientationally arbitrary but fixed mean velocity m_0 .

Lemma 4. Suppose $G(\hat{x})$ satisfies (H1)–(H4) and together with m_0 satisfies (H5), then $\dim(\text{Eig}(F_B^B, 0)) = 4$. Furthermore, all non-zero eigenvalues μ_i of F_B^B have negative real-part, $\mu_i \neq 0 \Rightarrow \text{Re}(\mu_i) < 0$.

Proof. Lemma 3 shows that there are no generalized eigenvectors for the zero eigenvalue of F_B^B . We will again skip the dependency of the matrix G on the stationary state. Let us first show that $\dim(\text{Eig}(F_B^B, 0)) \geq 4$. One directly checks that the basis elements $\{w_1, w_2, w_3\}$ of $\text{Eig}(G, 0)$ define linearly independent eigenvectors in $\text{Eig}(F_B^B, 0)$ by $z_i := (w_i, 0_{2N-2}, 0_2)^T$, $i = 1, 2, 3$. As remarked in Lemma 3, the lower 2×2 block of F_B^B can be diagonalized with eigenvalues 0 and -2α and associated eigenvectors $z_4 := (0_{2N}, 0_{2N-2}, m_0^\perp)^T$ and $(0_{2N}, 0_{2N-2}, m_0)^T$ respectively.

Next, we prove the last statement and $\dim(\text{Eig}(F_B^B, 0)) \leq 4$ simultaneously. First, note that any pair of eigenvalue μ and eigenvector $z = (x, v)^T$ satisfies $H z = \mu z$, which gives

$$v = \mu [x] \tag{18a}$$

$$\sum_{i=1}^N x_{2i-1} = \sum_{i=1}^N x_{2i} = 0 \tag{18b}$$

$$[Gx] - (I_{N-1} \otimes 2\beta(m_0 \otimes m_0^T))v = \mu v. \tag{18c}$$

Eq. (18c) rewrites as

$$\mu^2 [x] + (I_{N-1} \otimes 2\beta\mu(m_0 \otimes m_0^T))[x] = [Gx]. \tag{19}$$

By the same argument used in Lemma 3, (19) is equivalent to

$$\mu^2 x + (I_N \otimes 2\beta\mu(m_0 \otimes m_0^T))x = Gx, \tag{20}$$

due to (18b) and the properties of G . We left-multiply with \bar{x} to obtain

$$\mu^2 + 2\beta\mu \sum_{i=1}^N \left\langle m_0, x_{\frac{2i-1}{2i}} \right\rangle \overline{\left\langle m_0, x_{\frac{2i-1}{2i}} \right\rangle} = \bar{x}^T Gx, \tag{21}$$

assuming without loss of generality the normalization $\langle x, x \rangle = \bar{x}^T x = 1$ and denoting the complex scalar product as $\langle \cdot, \cdot \rangle$. Therefore, we deduce

$$\mu = -A \pm \sqrt{A^2 - B} \quad \text{with } A = -\beta \sum_{i=1}^N \left| \left\langle m_0, x_{\frac{2i-1}{2i}} \right\rangle \right|^2 \text{ and } B = -\bar{x}^T Gx.$$

Note that $A \geq 0, B \geq 0$ since $\bar{x}^T Gx \leq 0$ for any x due to (H3). We study three cases:

Case 1: Suppose $A > 0, B > 0$. If $A^2 - B \geq 0 \Rightarrow \mu < 0$ by monotonicity. If $A^2 - B < 0$, its square root is purely imaginary and thus $\text{Re}(\mu_i) < 0$.

Case 2: Suppose $A = 0, B > 0$. Then $\mu = \pm \sqrt{\bar{x}^T Gx} \in i\mathbb{R}$. Since A is a sum over squared moduli of N complex numbers, $A = 0 \Leftrightarrow \left\langle m_0, x_{\frac{2i-1}{2i}} \right\rangle = 0$ for all i . Since m is fixed, this holds if and only if x is particle-wise orthogonal to m_0 . Now, inserting $A = 0$ back into (20) gives $Gx = \mu^2 x, \mu^2 < 0$ which implies that such an x is also an eigenvector for G . This is a contradiction to (H5) and no such case exists.

Case 3: Suppose $A \geq 0, B = 0$. Then $x \in \text{Eig}(G, 0)$ and either $\mu = -2A < 0$ or $\mu = 0$. For the second case, $v = 0_{2N-2}$ follows from (18a), which implies $z \in \text{span}(z_1, z_2, z_3) \subset \text{Eig}(F_B^B, 0)$ by (H2).

Together, this shows that

$$\mu_i \neq 0 \Rightarrow \operatorname{Re}(\mu_i) < 0. \tag{22a}$$

$$A \geq 0, B = 0, \mu = 0 \Rightarrow z \in \operatorname{span}(z_1, z_2, z_3) \subset \operatorname{Eig}(F_B^B, 0). \tag{22b}$$

To complete the study of $\operatorname{Eig}(F_B^B, 0)$ suppose now that $\mu = 0$. Then $A^2 - B \geq 0$ and $0 = -A + \sqrt{A^2 - B} \Leftrightarrow A^2 = A^2 - B \Leftrightarrow B = 0$, which takes us back (22b). Therefore, no further eigenvector to eigenvalue zero exists, which completes the proof. ■

Let us remark that the 4 basis eigenvectors associated to the zero eigenvalue of F_B^B represent linearized flow within the set of stationary flock solutions \tilde{Z}_F in (11). From the change of variables (8) and Lemma 1, we know that the family of flock solutions is written as $\hat{Q} = (\hat{x}, 0_{2N-2}, m_0)^T$ with $|m_0|^2 = \frac{\alpha}{\beta}$. $G(\hat{x})$ is invariant under translation and spatial rotation, and thus z_1, z_2 , and z_3 are the linearized flow along these invariances. The fourth eigenvector z_4 represents linearized rotation of the mean velocity m_0 ...

Remark 1 (Comparison to the Techniques of [30]). The decisive technical difference used in the analysis presented here compared to [30] is the change of variables with respect to mean velocity. Here, we are able to show the absence of a generalized eigenvector in Lemmas 3 and 4. In the co-moving frame used in [30], this is not possible as a particular mean velocity is fixed, therefore the instability of a flock solution in the co-moving frame is shown. Instead, the instability criterion are in fact established proving the existence of a generalized eigenvectors associated to the zero eigenvalue (see [30, Remark 3.6]). The difference of the two approaches can also be seen in the proof of Lemma 4, which bears some technical similarity with [30, Lemma 3.1], but additionally uses the mean-velocity consistency in (18b) and (19), (20).

4. Proof of the main result

Changing coordinates to (8) eliminates the linear translation at constant speed and flock solutions become constant solutions which we identify with their profile and velocity at $t = 0$. The image of Z_F under (8) is \tilde{Z}_F in (11). \tilde{Z}_F is a manifold of stationary states of the flow of system (9). We aim to show that \tilde{Z}_F is a normally hyperbolic invariant manifold according to [33] for the time flow map of (9) restricted on $\operatorname{span}(B)$, which we denote by \mathbb{F}_B^t . First, \tilde{Z}_F is trivially invariant, $\mathbb{F}_B^t(\tilde{Z}_F) = \tilde{Z}_F$ for all $t \in \mathbb{R}$, as it is entirely stationary by Lemma 1. Choose $Q^* \in \tilde{Z}_F$ such that (H5) is fulfilled. Since eigenvectors are distinct, (H5) holds true for a neighbourhood of Q^* . There is no tangential flow along \tilde{Z}_F , and in particular

$$\frac{d}{dt} D_Q \mathbb{F}_B^t(Q^*) = F_B^B [D_Q \mathbb{F}_B^t(Q^*)]. \tag{23}$$

Lemmas 3 and 4 tell us, that F_B^B decomposes into the direct sum of $\operatorname{Eig}(F_B^B, 0)$, which equals the tangential space of \tilde{Z}_F at Q^* , and a subspace of eigenvectors whose eigenvalues have negative real part and span the non-tangential subspace of \tilde{Z}_F at Q^* . It is important to stress, that this splitting is preserved under the flow \mathbb{F}_B^t , since Q^* is stationary. Adapting from [33], all prerequisites for \tilde{Z}_F to be a normally hyperbolic invariant manifold are given. In the language of dynamical systems theory, the tangent bundle of \mathbb{R}^{4N} at Q^* has an \mathbb{F}_B -invariant splitting into tangential and stable sub-bundles. A necessary inequality between the tangential flow and the stable decay rate is trivially given in our case, since flow along \tilde{Z}_F is zero and as a consequence, we have a spectral gap. Note, that though \tilde{Z}_F is not compact we can easily select a compact sub-manifold around Q^* or perform a symmetry reduction and eliminate the translational invariance by further reducing $\operatorname{span}(B)$.

We can hence apply the theorem of Hadamard and Perron as in [33, Theorem 4.1], which, adapted to our situation, gives the following: There exists locally an \mathbb{F}_B -invariant stable smooth manifold in a neighbourhood of \tilde{Z}_F , where the splitting into stable and stationary of the linearized flow is preserved. The space between both manifolds is laminated by invariant fibres tangent to the stable subspace at Q^* , which are curves of the nonlinear flow towards \tilde{Z}_F . Hence, for a small perturbation of Q^* within that laminated space, we are sure to be sitting on some fibre, which might have Q^* as its base point on \tilde{Z}_F or a point in a neighbourhood of Q^* . In either case, there is transport along the fibre towards \tilde{Z}_F , for which an exponential decay estimate holds. This estimate is obtained by closing (23) on the stable sub-bundle. In other words, there exists locally a decomposition into fast and slow (stationary, in our case) coordinates of the nonlinear flow, such that the dynamics is characterized by the linearized flow. Going back to original coordinates (1), we obtain the statements in Theorem 1 and Corollary 1.

Remark 2. The precise meaning of “relaxation in time” towards a different flock solution in the statement of Theorem 1 is given by the notion of stable hyperbolicity of the family of flock solutions in coordinates (8), as used in the proof. In the original coordinates, there is no notion of “closeness” between elements of (7), since e.g. a change of orientation will cause two flocks to arbitrarily diverge in euclidean norm as $t \rightarrow \pm\infty$. Relaxation thus means, that the trajectory of a perturbed flock will approach a trajectory of another flock solution with typically different orientation and similar spatial configuration. Theorem 1 also tells us that the manifold is stable towards small perturbations of the flow \mathbb{F}_B , though that case is not considered here. Note, that this allows for flock stability under small perturbations of the potential.

Table 1

Numerical results for the largest non-zero eigenvalue μ_4 of the Jacobian $G(\{\hat{x}\})$ for varying potentials W and number of particles N . The numerical value of the zero eigenvalue μ_3 is shown as a measure of the exactness of $\{\hat{x}\}$. The free scaling parameter D is set to normalize the smallest eigenvalue to $\min(\mu_i) = -1$. As μ_4 is negative, hypotheses (H2)–(H3) of Theorem 1 are numerically validated for the cases shown.

	N	Eigenvalue structure of G		
		μ_4	$ \mu_3 $	D
Morse potential $C = \frac{10}{9}, \ell = \frac{3}{4}$	25	-6.9579e-5	5.3279e-9	9.5290
	40	-3.9449e-4	3.0435e-8	9.7423
	70	-2.3600e-4	6.8493e-9	10.0008
	100	-2.7069e-5	2.7073e-8	10.0720
Quasi-Morse potential $C = \frac{10}{9}, \ell = \frac{3}{4}, k = \frac{1}{2}$	25	-6.7374e-4	1.4273e-11	39.8647
	40	-1.6636e-4	3.5242e-12	39.0230
	70	-0.0024	9.1205e-12	38.5022
	100	-0.0023	4.0528e-11	38.3616
Generalized Morse potential $C = \frac{10}{9}, \ell = \frac{3}{4}, p = 1.25$	25	-5.0193e-5	4.5907e-8	7.6429
	40	-2.7459e-4	1.8389e-8	7.8653
	70	-1.2202e-4	1.7075e-10	7.9694
	100	-8.5882e-6	1.4491e-10	7.9813
Log-Newtonian potential	25	-3.6270e-4	4.0557e-6	0.25
	40	-1.8307e-8	8.0974e-9	0.25
	70	-0.0022	6.1172e-7	0.25
	100	-0.0029	7.1695e-6	0.25

5. Numerical experiments

In this section, we numerically investigate the hypothesis of Theorem 1 for a number of commonly used interaction potentials and study the nonlinear stability of flocks in terms of their polarization.

The key advantage of Theorem 1 is that essentially linear (asymptotic except symmetries) stability of the spatial configuration of a flock in the first-order aggregation model (2) is inherited nonlinearly by the family of flock solutions to the second-order model (1). In order to demonstrate that the stability theorem applies to a given potential, (H1)–(H4) have to hold for \hat{x} being a stationary state of (2). We compute a numerical approximation $\{\hat{x}\}$ of \hat{x} by a standard fourth-order Runge–Kutta time integration of (2) with random initial data and an interaction potential scaled with a factor D to relate the amplitude of the right-hand side of (2) with the accuracy of the solver:

$$W_D(r) := DW(r).$$

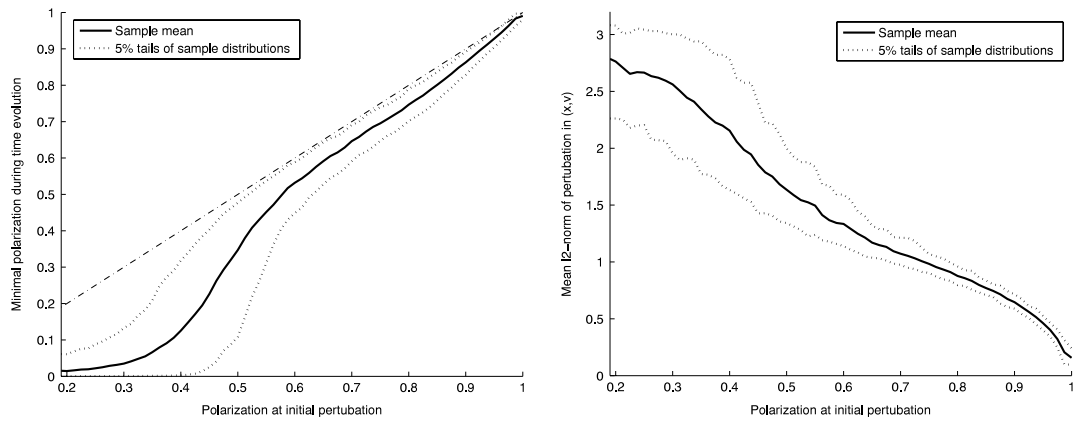
Note, that \hat{x} is independent of D . We hence obtain a spatial configuration for which (H1) holds numerically. Let us mention that (H4) is clearly satisfied for all potentials considered here. Hypotheses (H2)–(H3) require the Jacobian $G(\{\hat{x}\})$ to possess negative eigenvalues except for three zero eigenvalues associated to translation and rotation. As $G(\{\hat{x}\}) \in \mathbb{R}^{2N \times 2N}$ is symmetric it is sufficient to apply standard methods to compute its spectrum, at least for the number of particles we consider. Let $(\mu_1, \mu_2; \mu_3)$ denote the eigenvalues of $G(\{\hat{x}\})$ associated to translation and rotation and $\mu_i, i \geq 4$ all other eigenvalues sorted in descending order. In Table 1 we show μ_4 for different potentials and numbers of particles, together with $|\mu_3|$ to illustrate the numerical accuracy of $\{\hat{x}\}$ and the scaling parameter D . For the comparability of the potentials, D is set such that $\mu_{2N} = 1$ for all cases shown. The potentials used in Table 1 are:

	$W(r) = V(r) - CV(r/\ell)$
Morse	$V(r) = -e^{-r}$
Quasi-Morse	$V(r) = -\frac{1}{2\pi}K_0(kr)$
Generalized Morse	$V(r) = -e^{-\frac{r^p}{p}}$
Log-Newtonian	$W(r) = r^2 - \log(r)$

where K_0 is the modified Bessel function of second kind (for detailed discussions of the respective potentials see [32,31,12,34]).

It can be seen that for most cases considered, eigenvalues other than μ_1, μ_2, μ_3 are negative and hence Theorem 1 applies. The difference between the numerical eigenvalue μ_3 , which is equal to zero in theory, and μ_4 is of order four or higher, which shows that the negativity of μ_4 is not a numerical artefact. However, in some rare particular cases this conclusion cannot be drawn as μ_3 and μ_4 are of the same order (see the case $N = 40$, Log-Newtonian). We remark, that the eigenvalue computations of Table 1 require a highly accurate computation of $\{\hat{x}\}$, which is here obtained with a scaling of $D = 500$ and a time interval $T = [0, 500]$.

Next, we aim to study the stability of a flock in the nonlinear dynamics of (1) for the Morse potential under uniformly distributed perturbations in space and velocity. Theorem 1 provides a theoretical bound on the decay rate into equilibrium, which is very small as seen in Table 1. For random perturbations it is natural to assume that the flock is “more stable” than the theorem asserts. To study whether a flock remains a flock after an initial perturbation and the subsequent dynamics of



(a) Statistical comparison of initial polarization vs. minimal polarization in time evolution. (b) Statistics of the mean- ℓ_2 perturbation per particle in (x, v) for the initial perturbations of (a).

Fig. 1. Numerical findings of a sequence of simulations of system (1) with perturbed initial data. (a): Histogram of initial polarization of the perturbed flock against the corresponding minimal polarization observed during the subsequent time evolution. (b): Supporting statistics of the initial polarization against the average ℓ_2 -norm of perturbation per particle that induce the respective polarization change (—: arithmetic mean, - - : 5% tails of statistical distribution, ···: identity function, Morse potential as in Table 1, $\alpha = 1, \beta = 5, \lambda = 50, N = 100$ particles, 25,000 simulations).

(1), we use the notion of polarization of a group of individuals, which is defined as

$$\text{Pol}(\{(x_i, v_i)\}_{i=1}^N) := \frac{1}{N} \sum_i \frac{v_i}{\|v_i\|},$$

for N agents with phase-space coordinates at $(x_i, v_i) \in \mathbb{R}^4$, see e.g. [2] for instance. For an aligned flock, the polarization equals one whereas e.g. a rotating mill possesses zero polarization. We perform a sequence of simulations of (1) with randomly perturbed initial data. The stationary flock solution $z = \{(\hat{x}_i, m_0)\}_{i=1}^N, \|m_0\| = \frac{\alpha}{\beta}$ is perturbed by $\{(\Delta x_i, \Delta v_i)\}_{i=1}^N$ with $(\Delta x_i, \Delta v_i)$ sampled from the uniform distribution on $[-\frac{a}{2}, \frac{a}{2}] \times [-\frac{a}{2}, \frac{a}{2}]$ and varying perturbation strength a . Stability of the flock in the time evolution $[0, T]$ can be estimated as follows: If the support of the solution remains bounded and the polarization is equal or larger than the change in polarization induced by the initial perturbation, is it reasonable to assume that the ensemble will again tend to form a flock solution. If however, the polarization drops to significantly low or near zero values, the particles have evolved into an unpredictable state whose convergence behaviour as $T \rightarrow \infty$ is unknown. In Fig. 1(a), we plot the initial polarization induced by the perturbation against the minimal perturbation that occurred during a time integration of (1) with $T = 100$. The figure is based on 25,000 simulations with uniformly distributed perturbations strength a and shows the statistical mean and the 5% tails of the statistical distribution. In Fig. 1(b), we show the correlation between the polarization induced by the initial perturbation and the averaged ℓ_2 -norm of the perturbation per particle $\frac{1}{N} \sum_i \|(\Delta x_i, \Delta v_i)\|_2$, using the same statistical measures. The potential parameters are identical to Table 1 and the number of particles is $N = 100$. We observe from Fig. 1, that the probability of losing additional polarization in the flock nonlinearly increases with the strength of the initial perturbation. Below a critical threshold of approximately 0.6, the flock will most likely drop polarization with an increasing probability of near zero values (e.g. unpredictable state). From Fig. 1(b) we observe that a perturbation-induced polarization of 0.6 is on average generated by a strong average perturbation in (x, v) of approximately 1.34 per particle, which on average is a perturbation of approximately 0.95 in position per particle. However, the mean nearest neighbour distance of the stationary state considered here is approximately 0.252. Hence the nonlinear stability observed in this experiment well exceeds the validity regime of Theorem 1, which is expected to hold for spatial perturbations that are small compared to the nearest neighbour distances.

6. Conclusions & perspectives

In this paper, we have shown that flock solutions are fully nonlinearly stable under small enough position–velocity perturbations, in the sense that a perturbed flock configuration will relax to another flock configuration in $\tilde{\mathcal{Z}}_F$. Stability is inherited from the first-order particle system to the second-order model. Hence the flock might slightly rotate its position, translate its centre of mass, change the orientation of its mean velocity, or any combination of the three, but it will remain a flock solution. This result has been established for general interaction potentials W satisfying assumptions (H1)–(H4). We numerically investigated the spectrum of the Jacobian $G(\hat{x})$ evaluated at a stationary state \hat{x} for some commonly used potentials, and found them to possess the desired eigenvalue structure. Flocks for Morse-type potentials are linearly unstable in the co-moving frame second-order model [30], but nonlinearly asymptotically stable in the sense of our main result in Theorem 1. Finally, we investigated the nonlinear stability of flocks under uniformly distributed perturbations, and numerically validated the nonlinear stability of flocks in the sense of polarization.

Our findings are completely consistent with one's expectations from numerical experiments and, to the best of our knowledge, is the first of its kind for general flocks in second-order models. The task of theoretically proving nonlinear stability properties such as (H1)–(H4) of potentials and their steady solutions for the first-order aggregation equation remains largely open, and is inevitably complex given the extensive variety of stationary states (see e.g. [28]). An extension of the obtained result to three or higher dimensional space seems achievable, however with greater technical effort due to the increasing number of degrees of symmetry for the flock solution and the weakening of the interaction potential observed in [35]. The investigation of stability of other coherent patterns of swarming models such as rotating mills is another highly interesting open problem where the techniques used in this work clearly cannot be applied straightforwardly.

Acknowledgements

JAC acknowledges support from projects MTM2011-27739-C04-02, 2009-SGR-345 from Agència de Gestió d'Ajuts Universitaris i de Recerca-Generalitat de Catalunya, and the Royal Society through a Wolfson Research Merit Award. JAC, YH, and SM were supported by Engineering and Physical Sciences Research Council (UK) grant number EP/K008404/1. Supporting materials can be obtained from the corresponding author upon request.

References

- [1] J.K. Parrish, W.M. Hamner, *Animal Groups in Three Dimensions: How Species Aggregate*, Cambridge University Press, 1997.
- [2] I. Couzin, J. Krause, R. James, G.D. Ruxton, N.R. Franks, Collective memory and spatial sorting in animal group, *J. Theoret. Biol.* 218 (2002) 1–11.
- [3] I.D. Couzin, J. Krause, Self-organization and collective behavior of vertebrates, *Adv. Study Behav.* 32 (2003) 1–67.
- [4] F. Cucker, S. Smale, Emergent behavior in flocks, *IEEE Trans. Automat. Control* 52 (5) (2007) 852–862.
- [5] E. Cristiani, B. Piccoli, A. Tosin, Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints, in: *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Springer, 2010, pp. 337–364.
- [6] S. Camazine, J.-L. Deneubourg, N.R. Franks, J. Sneyd, G. Theraulaz, E. Bonabeau, *Self-organization in biological systems*, in: *Princeton Studies in Complexity*, Princeton University Press, Princeton, NJ, 2003, Reprint of the 2001 original.
- [7] A. Mogilner, L. Edelstein-Keshet, A non-local model for a swarm, *J. Math. Biol.* 38 (1999) 534–570.
- [8] D.D. Holm, V. Putkaradze, Aggregation of finite-size particles with variable mobility, *Phys. Rev. Lett.* 95 (2005) 226106.
- [9] D.D. Holm, V. Putkaradze, Formation of clumps and patches in selfaggregation of finite-size particles, *Physica D* 220 (2) (2006) 183–196.
- [10] H. Levine, W.-J. Rappel, I. Cohen, Self-organization in systems of self-propelled particles, *Phys. Rev. E* 63 (2000) 017101.
- [11] M.R. D'Orsogna, Y.-L. Chuang, A.L. Bertozzi, L.S. Chayes, Self-propelled particles with soft-core interactions: patterns, stability, and collapse, *Phys. Rev. Lett.* 96 (10) (2006) 104302.
- [12] Y.-L. Chuang, M.R. D'Orsogna, D. Marthaler, A.L. Bertozzi, L.S. Chayes, State transitions and the continuum limit for a 2D interacting, self-propelled particle system, *Physica D* 232 (1) (2007) 33–47.
- [13] J.A. Carrillo, M. Fornasier, G. Toscani, F. Vecil, Particle, kinetic, and hydrodynamic models of swarming, in: *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Springer, 2010, pp. 297–336.
- [14] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, O. Shochet, Novel type of phase transition in a system of self-driven particles, *Phys. Rev. Lett.* 75 (6) (1995) 1226.
- [15] P. Degond, S. Motsch, Continuum limit of self-driven particles with orientation interaction, *Math. Models Methods Appl. Sci.* 18 (supp01) (2008) 1193–1215.
- [16] J.A. Carrillo, M. Fornasier, J. Rosado, G. Toscani, Asymptotic flocking dynamics for the kinetic Cucker–Smale model, *SIAM J. Math. Anal.* 42 (1) (2010) 218–236.
- [17] I. Aoki, A simulation study on the schooling mechanism in fish, *Bull. Japan Soc. Sci. Fish* 48 (1982) 1081–1088.
- [18] A. Huth, C. Wissel, The simulation of fish schools in comparison with experimental data, *Ecol. Model.* 75/76 (1994) 135–145.
- [19] M. Ballerini, N. Cabbibo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, V. Zdravkovic, Interaction ruling animal collective behaviour depends on topological rather than metric distance: evidence from a field study, *Proc. Natl. Acad. Sci.* 105 (2008) 1232–1237.
- [20] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, R. Tavarone, From empirical data to inter-individual interactions: unveiling the rules of collective animal behavior, *Math. Models Methods Appl. Sci.* (2010) <http://dx.doi.org/10.1142/S0218202510004660>.
- [21] C.K. Hemelrijk, H. Hildenbrandt, Self-organized shape and frontal density of fish schools, *Ethology* 114 (2008) 245–254.
- [22] Alethea B.T. Barbaro, Kirk Taylor, Peterson F. Trethewey, Lamia Youseff, Björn Birnir, Discrete and continuous models of the dynamics of pelagic fish: application to the capelin, *Math. Comput. Simul.* 79 (12) (2009) 3397–3414.
- [23] H. Hildenbrandt, C. Carere, C.K. Hemelrijk, Self-organised complex aerial displays of thousands of starlings: a model, *Behav. Ecol.* 107 (21) (2010) 1349–1359.
- [24] T. Kolokolnikov, H. Sun, D. Uminsky, A. Bertozzi, Stability of ring patterns arising from 2d particle interactions, *Phys. Rev. E* 84 (1) (2011) 015203.
- [25] J.H. von Brecht, D. Uminsky, T. Kolokolnikov, A.L. Bertozzi, Predicting pattern formation in particle interactions, *Math. Models Methods Appl. Sci.* 22 (suppl. 1) (2012) 1140002, 31.
- [26] J.H. von Brecht, D. Uminsky, On soccer balls and linearized inverse statistical mechanics, *J. Nonlinear Sci.* 22 (6) (2012) 935–959.
- [27] D. Balagué, J.A. Carrillo, T. Laurent, G. Raoul, Nonlocal interactions by repulsive-attractive potentials: radial ins/stability, *Physica D* 260 (2013) 5–25.
- [28] D. Balagué, J.A. Carrillo, T. Laurent, G. Raoul, Dimensionality of local minimizers of the interaction energy, *Arch. Rat. Mech. Anal.* 209 (3) (2013) 1055–1088.
- [29] A.L. Bertozzi, J.H. von Brecht, H. Sun, T. Kolokolnikov, D. Uminsky, Ring patterns and their bifurcations in a nonlocal model of biological swarms, *Commun. Math. Sci.*, in preparation.
- [30] G. Albi, D. Balagué, J.A. Carrillo, J. von Brecht, Stability analysis of flock and mill rings for 2nd order models in swarming, 2013, Preprint.
- [31] J.A. Carrillo, S. Martin, V. Panferov, A new interaction potential for swarming models, *Physica D* 260 (2013) 112–126.
- [32] J.A. Carrillo, Y. Huang, S. Martin, Explicit flock solutions for quasi-morse potentials, 2013, Preprint.
- [33] M.W. Hirsch, C.C. Pugh, M. Shub, *Invariant Manifolds*, in: *Number no. 583 in Lecture Notes in Mathematics*, Springer-Verlag, 1977.
- [34] R.C. Fetecau, Y. Huang, T. Kolokolnikov, Swarm dynamics and equilibria for a nonlocal aggregation model, *Nonlinearity* 24 (10) (2011) 2681–2716.
- [35] F. Vecil, P. Lafitte, J. Rosado, A numerical study of attraction/repulsion collective behavior models: 3D particle analyses and 1D kinetic simulations, *Physica D* 260 (2013) 127–144.